ELSEVIER Theoretical Computer Science 290 (2003) 1629–1646

Theoretical Computer Science

www.elsevier.com/locate/tcs

On computing the entropy of cellular automata ☆

Michele D'amico[a], Giovanni Manzini[b], Luciano Margara[c,*]

[a] *Dipartimento di Matematica, Università di Bologna, Mura Anteo Zamboni 7 I-40127, Bologna, Italy*
[b] *Dipartimento di Scienze e Tecnologie Avanzate, Università del Piemonte Orientale "Amedeo Avogadro", Italy*
[c] *Dipartimento di Scienze dell'Informazione, Università di Bologna, Mura Anteo Zamboni 7 I-40127, Bologna, Italy*

Received 21 February 2001; received in revised form 10 October 2001; accepted 17 January 2002
Communicated by B. Durand

Abstract

We study the topological entropy of a particular class of dynamical systems: cellular automata. The topological entropy of a dynamical system (X,F) is a measure of the complexity of the dynamics of F over the space X. The problem of computing (or even approximating) the topological entropy of a given cellular automata is algorithmically undecidable (Ergodic Theory Dynamical Systems 12 (1992) 255). In this paper, we show how to compute the entropy of two important classes of cellular automata namely, *linear* and *positively expansive* cellular automata. In particular, we prove a closed formula for the topological entropy of D-dimensional ($D \geqslant 1$) linear cellular automata over the ring $\mathbf{Z}_m$ ($m \geqslant 2$) and we provide an algorithm for computing the topological entropy of positively expansive cellular automata.

1. Introduction

Cellular automata (CA) are dynamical systems consisting of a regular lattice of variables which can take a finite number of discrete values. The global state of the CA, specified by the values of all the variables at a given time, evolves according to a global transition map F based on a *local rule* f which acts on the value of each single cell in synchronous discrete time steps.

☆ A preliminary version of this paper has been presented to the 25th International Colloquium on Automata, Languages, and Programming (ICALP '98).

* Corresponding author.
E-mail addresses: damico@cs.unibo.it (M. D'amico), manzini@mfn.unipmn.it (G. Manzini), margara@cs.unibo.it (L. Margara).

0304-3975/02/$ - see front matter
PII: S0304-3975(02)00071-3

A CA can be viewed as a discrete time dynamical system (X,F) where $F:X\to X$ is the CA global transition map defined over the configuration space X. The *dynamical behavior* of CA can be analyzed—as that of any other dynamical system—in different frameworks. For example, in [2] the authors study measure theoretic properties of CA, while in [3,11,12] the authors investigate the topological behavior of CA. The classical problem in CA theory is the following: given a description of the local rule f, determine whether the global transition map F associated to f satisfies a certain property. In the case of general CA, this problem is algorithmically unsolvable for a number of important properties, e.g., surjectivity and injectivity are undecidable in any dimension greater than one [8], nilpotency is undecidable also for 1-dimensional CA [7], topological entropy of 1-dimensional CA is not even approximable [5]. Finally, it is a common belief that also dynamical properties such as sensitivity, equicontinuity, transitivity, and ergodicity are undecidable even if, to our knowledge, no formal proof of this fact has been produced so far. On the other hand, if we restrict to particular subclasses of CA, many of the above properties become decidable (often in polynomial time). For example, injectivity and surjectivity are decidable for 1-dimensional CA [1] and all the above mentioned dynamical properties are decidable for D-dimensional linear CA over $\mathbf{Z}_m$ [3,6,10,11,12].

In this paper, we focus our attention on topological entropy which is one of the most studied properties of dynamical systems [13,14,18]. Informally, the topological entropy measures the *uncertainty* of the forward evolution of any dynamical system in the presence of incomplete description of initial configurations. Since CA are deterministic dynamical systems, given a complete description of any configuration $x\in X$ we may determine exactly its forward trajectory $T=\{x,F(x),F^2(x),\ldots\}$ according to F. Topological entropy gives a quantitative estimation of the uncertainty we introduce in T assuming we do not have a complete description of x.

The topological entropy of general CA cannot be algorithmically computed [5]. Nevertheless it is important to investigate for which classes of CA topological entropy can be computed and how to accomplish this task. The main contribution of this paper is the solution to the two following open problems addressed in [5] and [2], respectively.

Problem 1. In [5] the authors prove the undecidability of topological entropy and conclude that "… *the undecidability question remains open if one restricts to a subclass of cellular automata such as linear rules* …". In Theorems 2 and 3, we prove a closed formula for the topological entropy of D-dimensional linear CA over $\mathbf{Z}_m$ (in terms of the coefficients of the local rule associated to the CA) for $D=1$ and for $D\geqslant 2$, respectively. Note that Theorem 3 is a refinement of a recent result by Morris and Ward [14] which states that the entropy of a linear D-dimensional CA with $D\geqslant 2$, is either zero or infinity.

Problem 2. In [2] the authors review topological and metric properties of CA and prove that "… *the topological entropies of positively expansive cellular automata are log-integers* …" leaving open the problem of computing the entropy for that class of CA. In Theorems 5 and 6, we show how to efficiently compute the entropy of positively expansive CA.

We also give a closed formula for the Lyapunov exponents of 1-dimensional linear CA over $\mathbf{Z}_m$ (Theorem 1). This result, together with the solution of Problem 1, partially answers a question arising at page in [17, p. 9] where the authors ask for classes of CA for which the topological entropy is equal to the sum of the left and the right Lyapunov exponents multiplied by the entropy of the shift CA. It turns out (Corollary 4) that, as far as linear CA are concerned, topological entropy satisfies the above-mentioned property if and only if the local rule associated to the CA is leftmost and rightmost permutive (the notion of permutivity was first introduced by Hedlund in [4]).

The rest of this paper is organized as follows. In Section 2, we give basic definitions and notations. In Section 3, we state our results and we give detailed proofs in Section 3. Section 5 contains some concluding remarks.

2. Basic definitions

In this section, we review some basic definitions and state some known results concerning CA.

2.1. Cellular automata

For $m \geqslant 2$, let $\mathbf{Z}_m = \{0, 1, \ldots, m-1\}$. We consider the *space of configurations*

$$\mathscr{C}_m^D = \{c \mid c : \mathbf{Z}^D \to \mathbf{Z}_m\}$$

which consists of all functions from $\mathbf{Z}^D$ into $\mathbf{Z}_m$. Each element of $\mathscr{C}_m^D$ can be visualized as an infinite D-dimensional lattice in which each cell contains an element of $\mathbf{Z}_m$. Let $s \geqslant 1$. A *neighborhood frame* of size s is an ordered set of distinct vectors $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_s \in \mathbf{Z}^D$. Given $f : \mathbf{Z}_m^s \to \mathbf{Z}_m$, a D-dimensional CA based on the *local rule* f is the pair $(\mathscr{C}_m^D, F)$, where $F : \mathscr{C}_m^D \to \mathscr{C}_m^D$, is the *global transition map* defined as follows.

$$[F(c)](\vec{v}) = f(c(\vec{v} + \vec{u}_1), \ldots, c(\vec{v} + \vec{u}_s)), \quad \text{where } c \in \mathscr{C}_m^D,\ \vec{v} \in \mathbf{Z}^D. \tag{1}$$

In other words, the content of cell $\vec{v}$ in the configuration $F(c)$ is a function of the content of cells $\vec{v} + \vec{u}_1, \ldots, \vec{v} + \vec{u}_s$ in the configuration c. Note that the local rule f and the neighborhood frame completely determine F.

For 1-dimensional CA we use a simplified notation. A local rule $f : \mathbf{Z}_m^{2r+1} \to \mathbf{Z}_m$ of *radius* r is denoted by $f(x_{-r}, \ldots, x_{-1}, x_0, x_1, \ldots, x_r)$. The associated global map $F : \mathscr{C}_m^1 \to \mathscr{C}_m^1$ is defined by

$$[F(c)](i) = f(x_{-r+i}, \ldots, x_{r+i}), \quad \text{where } c \in \mathscr{C}_m^1,\ i \in \mathbf{Z}.$$

We assume that f explicitly depends on at least one of the two variables x_{-r} and x_r. We say that f is *permutive* in the variable x_i, $-r \leqslant i \leqslant r$, if and only if, no matter which values are given to the other $2r$ variables, the modification of the value of x_i causes the modification of the output produced by f (for a formal definition of permutivity see Definition 6 of [4]).

Throughout the paper, $F(c)$ will denote the result of the application of the map F to the configuration c, and $c(\vec{v})$ will denote the value assumed by c in $\vec{v}$. For $n \geqslant 0$,

we recursively define $F^n(c)$ by $F^n(c)=F(F^{n-1}(c))$, where $F^0(c)=c$. Let $(\mathscr{C}_m^D,F)$ be a CA based on the local rule f. We denote by $f^{(n)}$ the local rule associated to F^n.

2.2. Linear CA over $\mathbf{Z}_m$

In the special case of linear CA the set $\mathbf{Z}_m$ is endowed with the usual sum and product operations that make it a commutative ring. In what follows we denote by $[x]_m$ the integer x taken modulo m. Linear CA have a local rule of the form $f(x_1,\ldots,x_s)=[\sum_{i=1}^{s}\lambda_i x_i]_m$ with $\lambda_1,\ldots,\lambda_s\in\mathbf{Z}_m$. Hence, for a linear D-dimensional CA (1) becomes

$$[F(c)](\vec{v})=\left[\sum_{i=1}^{s}\lambda_i c(\vec{v}+\vec{u}_i)\right]_m,\quad \text{where } c\in\mathscr{C}_m^D,\ \vec{v}\in\mathbf{Z}^D. \tag{2}$$

For linear 1-dimensional CA the local rule f can be written as $f(x_{-r},\ldots,x_r)=[\sum_{i=-r}^{r}a_i x_i]_m$ where at least one between a_{-r} and a_r is nonzero. In this case (1) becomes

$$[F(c)](i)=\left[\sum_{j=-r}^{r}a_j c(i+j)\right]_m,\quad \text{where } c\in\mathscr{C}_m^1,\ i\in\mathbf{Z}.$$

Note that f is permutive in the jth variable iff $\gcd(a_j,m)=1$.

In the following we will make use also of the *formal power series* (fps) representation of the configuration space $\mathscr{C}_m^D$ (see [6, Section 3] for details). For example, for $D=1$, to each configuration $c\in\mathscr{C}_m^1$ we associate the fps $P_c(X)=\sum_{i\in\mathbf{Z}}c(i)X^i$. The advantage of this representation is that the computation of a linear map is equivalent to power series multiplication. Let $F:\mathscr{C}_m^1\to\mathscr{C}_m^1$ be a linear map with local rule $f(x_{-r},\ldots,x_r)=[\sum_{i=-r}^{r}a_i x_i]_m$. We associate to F the finite fps $A_f(X)=\sum_{i=-r}^{r}a_i X^{-i}$. Then, for any $c\in\mathscr{C}_m^1$ we have

$$P_{F(c)}(X)\equiv P_c(X)A_f(X)\,(\mathrm{mod}\,m). \tag{3}$$

Note that each coefficient of $P_{F(c)}(X)$ is well defined since $A_f(X)$ has only finitely many non-zero coefficients. Note also that the finite fps associated to F^n is $A_f^n(X)$. In the D-dimensional case to each configuration $c\in\mathscr{C}_m^D$ we associate the formal power series

$$P_c(X_1,\ldots,X_D)=\sum_{i_1,\ldots,i_D\in\mathbf{Z}}c(i_1,\ldots,i_D)X_1^{i_1}\cdots X_D^{i_D}.$$

The computation of a linear map F over $\mathscr{C}_m^D$ is equivalent to the multiplication by a finite fps $A(X_1,\ldots,X_D)$ which is defined in terms the local rule f and the neighborhood frame $\vec{u}_1,\ldots,\vec{u}_s$. The finite fps associated to the map F defined by (2) is

$$A(X_1,\ldots,X_D)=\sum_{i=1}^{s}\lambda_i X_1^{-\vec{u}_i(1)}\cdots X_D^{-\vec{u}_i(D)},$$

where $\vec{u}_i(j)$ denotes the jth component of vector $\vec{u}_i$.

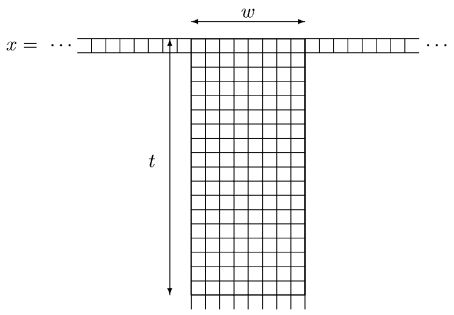

Fig. 1. $R(w,t)$ is the number of distinct rectangles (distinct "snapshots" of the cells inside the thick lines) we obtain by iterating F on all different configurations x.

2.3. Topological entropy of CA

The topological properties of CA are usually defined with respect to the metric topology induced by the *Tychonoff distance* over the configuration space $\mathscr{C}_m^D$ (see for example [2]). With this topology $\mathscr{C}_m^D$ is a Cantor set i.e. a compact, perfect and totally disconnected set, and every CA is a (uniformly) continuous map. The definition of topological entropy $\mathscr{H}$ of a continuous map $F:X\to X$ over a compact space X can be found for example in [2]. The value $\mathscr{H}(X,F)$ is generally accepted as a measure of the complexity of the dynamics of F over X. In [5] it is shown that for 1-dimensional CA the general definition of topological entropy translates to the following simpler form. Let $R(w,t)$ denote the number of distinct rectangles of width w and height t occurring in a space-time evolution diagram of $(\mathscr{C}_m^1,F)$ (see Fig. 1). Let r denote the radius of the local rule f associated to F, it is easy to see that

$$m^w \leqslant R(w,t) \leqslant m^{w+2r(t-1)}.$$

Given w and t, $R(w,t)$ can be determined by computing the evolution of all blocks of length $w+2r(t-1)$. The topological entropy of F is given by

$$\mathscr{H}(\mathscr{C}_m^1,F) = \lim_{w\to\infty}\left(\lim_{t\to\infty}\frac{\log R(w,t)}{t}\right). \tag{4}$$

From (4) it follows that the entropy of a 1-dimensional CA over $\mathscr{C}_m^1$ satisfies $\mathscr{H}(F)\leqslant 2r\log m$.

Example 1. An important 1-dimensional CA is the *right shift map* $(\mathscr{C}_m^1,\sigma)$ defined by $[\sigma(c)](i)=c(i-1)$. The global map σ corresponds to the local rule $f(x_{-1},x_0,x_1)=x_{-1}$. Given $w,t>0$ we have $R(w,t)=m^{w+t-1}$, hence, $\mathscr{H}(\mathscr{C}_m^1,\sigma)=\log m$.

For D-dimensional CA, we can still compute $\mathscr{H}$ using (4) provided $R(w,t)$ is replaced by $R(w^{(D)},t)$ which denotes the number of distinct $(D+1)$-dimensional hyperrectangles obtained as space-time evolution diagrams of $(\mathscr{C}_m^D,F)$. Now, w has to be interpreted as the side-length of a D-dimensional region of the lattice.

2.4. Lyapunov exponents for CA

We recall the definition of Lyapunov exponents for the special case of 1-dimensional CA given in [17]. There, the authors introduce quantities analogous to Lyapunov exponents of smooth dynamical systems with the aim of describing the *local instability* of orbits in CA. For every $x \in \mathscr{C}_m^1$ and $s \geqslant 0$ we set

$$W_s^+(x) = \{y \in \mathscr{C}_m^1 : \ y(i) = x(i) \text{ for all } i \geqslant s\},$$

$$W_s^-(x) = \{y \in \mathscr{C}_m^1 : \ y(i) = x(i) \text{ for all } i \leqslant -s\}.$$

We have that $W_i^+(x) \subset W_{i+1}^+(x)$ and $W_i^-(x) \subset W_{i+1}^-(x)$. For every $n \geqslant 0$ we define

$$\tilde{\Lambda}_n^+(x) = \min\{s \geqslant 0 : \ F^n(W_0^+(x)) \subset W_s^+(F^n(x))\},$$

$$\tilde{\Lambda}_n^-(x) = \min\{s \geqslant 0 : \ F^n(W_0^-(x)) \subset W_s^-(F^n(x))\}.$$

The above definitions may appear a little confusing, but the values $\tilde{\Lambda}_n^+(x)$ and $\tilde{\Lambda}_n^-(x)$ have a simple intuitive meaning. $W_0^+(x)$ is the set of configurations which agree with x in all cells with index $i \geqslant 0$. By comparing $F^n(x)$ with $F^n(W_0^+(x))$ the value $\tilde{\Lambda}_n^+(x)$ measures how far differences in cells with index $i < 0$ can "propagate" to the right-hand side (that is, towards the cells with positive index) in n iterations of F. Similarly, $\tilde{\Lambda}_n^-(x)$ measures how far differences in cells with index $i > 0$ can "propagate" to the left-hand side in n iterations of F. For example, for the right shift map σ defined in Example 1, we have $\tilde{\Lambda}_n^+(x) = n$ and $\tilde{\Lambda}_n^-(x) = 0$ for all $x \in \mathscr{C}_m^1$.

We also introduce the following shift invariant quantities:

$$\Lambda_n^-(x) = \max_{j \in \mathbf{Z}} \tilde{\Lambda}_n^-(\sigma^j(x)), \quad \Lambda_n^+(x) = \max_{j \in \mathbf{Z}} \tilde{\Lambda}_n^+(\sigma^j(x)), \tag{5}$$

where σ denotes the right shift map. Intuitively, the value $\tilde{\Lambda}_n^+(\sigma^j(x))$ (resp. $\tilde{\Lambda}_n^-(\sigma^j(x))$) measures how far differences in cells with index $i < j$ (resp. $i > j$) can "propagate" to the right (resp. left) in n iterations of F. Finally, the values $\lambda^+(x)$ and $\lambda^-(x)$ defined by

$$\lambda^+(x) = \lim_{n \to \infty} \frac{1}{n} \Lambda_n^+(x), \quad \lambda^-(x) = \lim_{n \to \infty} \frac{1}{n} \Lambda_n^-(x) \tag{6}$$

are called, respectively, the right and left Lyapunov exponents of the CA F for the configuration x. If F is linear it is easy to see that $\lambda^+(x)$ and $\lambda^-(x)$ do not depend on x, i.e., there exist two constants λ^+ and λ^- such that for every $x \in \mathscr{C}_m^1$, $\lambda^-(x) = \lambda^-$ and $\lambda^+(x) = \lambda^+$. The following result is a simple corollary of Theorem 2 of [17, p. 5]. For any CA F we have

$$\mathscr{H}(\mathscr{C}_m^1, F) \leqslant \mathscr{H}(\mathscr{C}_m^1, \sigma)(\lambda^+ + \lambda^-), \tag{7}$$

where $\mathscr{H}(\mathscr{C}_m^1, \sigma) = \log m$ is the entropy of the shift map (see Example 1).

2.5. Dynamical properties of CA

For any discrete time dynamical system defined on a metric space, it is possible to define important properties which provide useful information on the long term behavior of the system. We now recall the definition of some of these properties for a generic system (X,F), where $F:X\to X$. Here, we assume that X is equipped with a distance d and that the map F is continuous on X according to the metric topology induced by d (for CA, the Tychonoff distance satisfies this property). We denote by $\mathscr{B}(x,\varepsilon)$ the (open) set $\{y\in X\colon d(x,y)<\varepsilon\}$.

Definition 1 (*Positive expansivity*). A discrete time dynamical system (X,F) is positively expansive if and only if there exists $\delta>0$ such that for every $x,y\in X,\ x\neq y$ there exists $n\geqslant 0$ such that $d(F^n(x),F^n(y))>\delta$. The value δ is called the expansivity constant.

Intuitively, a map is positively expansive if every pair of enough close points eventually separate by at least δ under iteration of F. If a map is positively expansive, then, for all practical purposes, the dynamics of the map defies numerical approximation. Small errors in computation which are introduced by round-off become magnified upon iteration.

Definition 2 (*Sensitivity*). A discrete time dynamical system (X,F) is sensitive to initial conditions if and only if there exists $\delta>0$ such that for any $x\in X$ and for any $\varepsilon>0$, there exist $y\in\mathscr{B}(x,\varepsilon)$ and $n\geqslant 0$, such that $d(F^n(x),F^n(y))>\delta$. The value δ is called the sensitivity constant.

Intuitively, a map is sensitive to initial conditions, or simply sensitive, if there exist points arbitrarily close to x which eventually separate from x by at least δ under iteration of F. Note that not all points near x need eventually separate from x under iteration, but there must be at least one such point in every neighborhood of x.

Definition 3 (*Equicontinuity at x*). A discrete time dynamical system (X,F) is equicontinuous at $x\in X$ if and only if for any $\delta>0$ there exists $\varepsilon>0$ such that for any $y\in\mathscr{B}(x,\varepsilon)$ and $n\geqslant 0$ we have $d(F^n(x),F^n(y))<\delta$.

Definition 4 (*Equicontinuity*). A discrete time dynamical system (X,F) is equicontinuous if and only if it is equicontinuous at every $x\in X$.

The notions of sensitivity and equicontinuity are related. In fact, by comparing the definitions one can easily see that

$$F \text{ is not sensitive } \Leftrightarrow\ \exists x\colon F \text{ is equicontinuous at } x. \tag{8}$$

For linear CA we have that $(\mathscr{C}_m^D, F)$ is equicontinuous if and only if it is equicontinuous in a single point $x \in \mathscr{C}_m^D$. Hence, (8) becomes

$$(\mathscr{C}_m^D, F) \text{ is not sensitive } \Leftrightarrow \ (\mathscr{C}_m^D, F) \text{ is equicontinuous.} \tag{9}$$

Sensitive linear CA, hence, in view of (9) equicontinuous CA, are completely characterized by the following theorem (see [12, Theorem 6]).

Theorem 5. *Let F denote the global transition map of a linear D-dimensional CA over $\mathbf{Z}_m$ defined by*

$$[F(c)](\vec{v}) = \left[\sum_{i=1}^{s} \lambda_i c(\vec{v} + \vec{u}_i)\right]_m .$$

Assume $\vec{u}_1 = \vec{0}$, that is, λ_1 is the coefficient associated to the null displacement. Then F is sensitive if and only if there exists a prime p such that

$$p \mid m \quad \text{and} \quad p \nmid \gcd(\lambda_2, \lambda_3, \ldots, \lambda_s).$$

In other words, F is sensitive unless every prime which divides m divides also all the coefficients λ_i's corresponding to nonzero neighborhood vectors.

3. Statement of new results

In this section we state our main results. In particular, Section 3.1 contains results concerning linear CA, while Section 3.2 contains results about positively expansive CA.

3.1. Entropy of linear CA

Our first result provides a closed formula for the Lyapunov exponents of 1-dimensional linear CA.

Theorem 1. *Let $(\mathscr{C}_m^1, F)$ be a 1-dimensional CA over $\mathbf{Z}_m$ with local rule $f(x_{-r}, \ldots x_r) = [\sum_{i=-r}^{r} a_i x_i]_m$, and let $m = p_1^{k_1} \cdots p_h^{k_h}$ be the prime factor decomposition of m. For $i = 1, \ldots, h$ define*

$$P_i = \{0\} \cup \{j : \gcd(a_j, p_i) = 1\}, \quad L_i = \min P_i, \qquad R_i = \max P_i.$$

Then, the right and left Lyapunov exponents of $(\mathscr{C}_m^1, F)$ are

$$\lambda^+ = -\min_{1 \leqslant i \leqslant h} \{L_i\} \quad \text{and} \quad \lambda^- = \max_{1 \leqslant i \leqslant h} \{R_i\}.$$

In the next theorem we give a closed formula for the entropy of 1-dimensional linear CA which can be efficiently computed in terms of the coefficients of the local rule.

Theorem 2. *Let* $(\mathscr{C}_m^1, F)$ *be a* 1*-dimensional CA over* $\mathbf{Z}_m$ *with local rule* $f(x_{-r},\dots x_r)=[\sum_{i=-r}^{r} a_i x_i]_m$, *and let* $m=p_1^{k_1}\cdots p_h^{k_h}$ *denote the prime factor decomposition of* m. *Let* L_i *and* R_i *be defined as in Theorem* 3.1. *Then*

$$\mathscr{H}(\mathscr{C}_m^1, F) = \sum_{i=1}^{h} k_i(R_i - L_i)\log(p_i). \tag{10}$$

In the next example we use the above theorems to compute the entropy and the Lyapunov exponents of a 1-dimensional linear CA.

Example 2. For $m=1620=2^2 3^4 5$, consider the linear local rule

$$f(x_{-2},\dots,x_4) = [(10x_{-2} + 15x_{-1} + 9x_0 + 18x_1 + 22x_2 + 4x_3 + 30x_4)]_{2^2 3^4 5},$$

and let $(\mathscr{C}_{1620}^1, F)$ be the 1-dimensional linear CA associated with f. From Theorem 2 we have

$$L_1 = -1, \quad L_2 = -2, \quad L_3 = 0, \quad R_1 = 0, \quad R_2 = 3, \quad R_3 = 3,$$

and then $\mathscr{H}(\mathscr{C}_{1620}^1, F)=2+20\log 3+4\log 5$. In addition, according to Theorem 1 we have $\lambda^+=2$ and $\lambda^-=3$.

In [14], the authors prove that for D-dimensional linear CA with $D\geqslant 2$ the topological entropy must be 0 or infinity. In the next theorem, using the classification given in [12], we prove that it is possible to easily check if a D-dimensional linear CA has zero or infinite entropy.

Theorem 3. *Let* $(\mathscr{C}_m^D, F)$ *be a* D*-dimensional linear CA over* $\mathbf{Z}_m$ *with* $D\geqslant 2$. *Then, either* F *is sensitive to initial conditions and* $\mathscr{H}(\mathscr{C}_m^D, F)=\infty$, *or* F *is equicontinuous and* $\mathscr{H}(\mathscr{C}_m^D, F)=0$.

3.2. Entropy of positively expansive CA

Since positively expansive CA do not exist in any dimension greater than 1 (see [18]), in this section we deal only with 1-dimensional CA. However, in our proofs we only use the fact that each positively expansive CA is a surjective open map (see [9]) and is topologically conjugated to a one-sided full shift (see [1] Theorem 8.5 in [15]).

We now introduce some notation we need in order to state the result of this section. To any CA F based on the local rule f (with radius r) we associate a directed labeled graph $G_F=(V,E)$ called *finite predecessor graph* (fp-graph) defined as follows. The vertex set $V=\{s \mid s\in \mathbf{Z}_m^{2r}\}$ is the set of all strings of length $2r$ from the alphabet $\mathbf{Z}_m$. The edge set E consists of all ordered pairs $(s_1,s_2)\in V\times V$ such that $s_1(i+1)=s_2(i)$, for $i=1,\dots,2r-1$. The label a associated to (s_1,s_2) is $a=f(s_1(1),\dots,s_1(2r),s_2(2r))$. As usual, we write $(s_1 \xrightarrow{a} s_2)$ to denote such edge.

[1] Since [15] is a rather hard to find reference, the reader may obtain an alternative proof of this statement by combining Theorem 3.12 of [16] and Theorem 10 of [9].

Example 3. *Let* $f:\{0,1\}^3\to\{0,1\}$ be the radius 1 local rule defined by $f(x_{-1},x_0,x_1)=[(x_{-1}+x_1)]_2$. Let F denote the global transition map associated to f. By Theorem 7 in [12] F is expansive. The fp-graph $G_F=(V,E)$ has vertex set $V=\{00,01,10,11\}$, and edge set

$$E=\{(00\xrightarrow{0}00),(00\xrightarrow{1}01),(01\xrightarrow{0}10),(01\xrightarrow{1}11),(10\xrightarrow{1}00),(10\xrightarrow{0}01),\\ \times\ (11\xrightarrow{0}11),(11\xrightarrow{1}10)\}.$$

Let f denote the local rule of a one-dimensional CA of radius r. We know that f is a map from $\mathbf{Z}_m^{2r+1}$ to $\mathbf{Z}_m$. For any $\ell\geqslant 1$ we define the map $f_\ell:\mathbf{Z}_m^{2r+\ell}\to\mathbf{Z}_m^{\ell}$ as follows. Given $\alpha\in\mathbf{Z}_m^{2r+\ell}$ we obtain $f_\ell(\alpha)\in\mathbf{Z}_m^{\ell}$ applying f to the $(2r+1)$-tuples of α (there are exactly ℓ of them). In other words, the ith component of $f_\ell(\alpha)$ is $f(\alpha_{i-r},\ldots,\alpha_i,\alpha_{i+r})$.

Definition 4. Let F be a CA with local rule f. If $\alpha\in\mathbf{Z}_m^{2r+\ell}$, $\beta\in\mathbf{Z}_m^{\ell}$ and $\beta=f_\ell(\alpha)$ we say that α is a *finite predecessor* of β according to F.

There is a strict relation between the fp-graph G_F and the concept of finite predecessor. Given any path $s_0\xrightarrow{a_1}s_1\xrightarrow{a_2}\cdots\xrightarrow{a_\ell}s_\ell$ of G_F we say that the sequence $s=s_0,\ldots,s_\ell$ is labeled by $a_1,\ldots,a_\ell$. From the definition of G_F we have that $s_0(1),\ldots,s_0(2r),s_1(2r),\ldots,s_l(2r)$ is a finite predecessor of $a_1,\ldots,a_\ell$ that is

$$f_\ell(s_0(1),\ldots,s_0(2r),s_1(2r),\ldots,s_\ell(2r))=a_1,\ldots,a_\ell.$$

It is straightforward to prove that G_F has the following property: the cardinality of the set of finite predecessors of any given sequence s of length ℓ is given by the number of distinct paths in G_F labeled by s.

Given a length-n string $\alpha\in(\mathbf{Z}_m)^n$ we denote by $\bar{\alpha}$ the 1-dimensional configuration consisting of the concatenation of infinitely many copies of α. Formally, for any $i\in\mathbf{Z}$, $\bar{\alpha}(i)=\alpha([i]_n)$.

The next theorem shows that using the fp-graph of a positively expansive CA $(\mathscr{C}_m^1,F)$ we can compute the number of predecessors of any $c\in\mathscr{C}_m^1$, that is, the cardinality of the set $F^{-1}(c)$.

Theorem 5. *Let* $(\mathscr{C}_m^1,F)$ *be a positively expansive CA. Then* F *is a* μ*-to-one map, that is, every configuration* $c\in\mathscr{C}_m^1$ *has exactly* μ *predecessors. The value* μ *is the sum of the lengths of all distinct cycles of* $G_F=(V,E)$ *labeled by a finite sequence of* 0*'s.*

Example 4. Consider the map F defined in Example 3. In the graph G_F there are 3 cycles labeled by sequences of type 0^k namely

$$00\xrightarrow{0}00,\quad 01\xrightarrow{0}10\xrightarrow{0}01,\quad 11\xrightarrow{0}11.$$

We conclude that the number of predecessors of the configuration $\bar{0}$ is $1+2+1=4$. Indeed, they are $\bar{0}$, $\overline{10}$, $\overline{01}$, and $\bar{1}$. Since F is a μ-to-one mapping every other configuration has the same number of predecessors.

The following theorem, combined with Theorem 5, shows how to compute the topological entropy of any positively expansive CA.

Theorem 6. *The entropy of any positively expansive CA* $(\mathscr{C}_m^1, F)$ *is equal to* $\log \mu$, *where* μ *is the number of predecessors of any configuration* $c \in \mathscr{C}_m^1$ *and is given by Theorem* 3.5.

4. Proof of the main theorems

This section contains the proofs of the results stated in Section 3.

4.1. Entropy of linear CA

Let $(\mathscr{C}_m^D, F)$ be a linear CA, and let q be any factor of m. For any configuration $c \in \mathscr{C}_m^D$, $[c]_q$ will denote the configuration in $\mathscr{C}_q^D$ defined by

$$[c]_q(\vec{v}) = [c(\vec{v})]_q = c(\vec{v}) \bmod q \quad \text{for all } \vec{v} \in \mathbf{Z}^D.$$

Similarly, $[F]_q$ will denote the map $[F]_q : \mathscr{C}_q^D \to \mathscr{C}_q^D$ defined by $[F]_q(c) = [F(c)]_q$.

The following lemma shows that for any 1-dimensional linear CA $(\mathscr{C}_{p^k}^1, F)$ with p prime which is not nilpotent there exists $h \geqslant 1$ such that F^h is permutive in both its *leftmost* and *rightmost* variables (permutivity has been defined at the end of Section 2).

Lemma 1. *Let* $(\mathscr{C}_{p^k}^1, F)$ *be a linear* 1*-dimensional CA over* $\mathbf{Z}_{p^k}$ (p *prime*) *with local rule* $f(x_1, \ldots, x_s) = [\sum_{i=1}^{s} a_i x_i]_{p^k}$. *Assume there exists* a_l *such that* $\gcd(a_l, p) = 1$, *and let*

$$\hat{P} = \{j:\ \gcd(a_j, p) = 1\}, \quad \hat{L} = \min \hat{P}, \quad \hat{R} = \max \hat{P}.$$

Then, there exists $h \geqslant 1$ *such that the local rule* $f^{(h)}$ *associated to* F^h *has the form*

$$f^{(h)}(x_{-hr}, \ldots, x_{hr}) = \left[\sum_{i=h\hat{L}}^{h\hat{R}} b_i x_i \right]_{p^k} \quad \text{with} \quad \gcd(b_{h\hat{L}}, p) = \gcd(b_{h\hat{R}}, p) = 1. \tag{11}$$

Proof. We associate to the local rule f the formal power series $A(X) = \sum_{i=-r}^{r} a_i X^{-i}$ defined in Section 2.2. As we have already pointed out, the formal power series associated to $f^{(n)}$ is $A^n(X)$. Let $A(X) = A_1(X) + pA_2(X)$, where $A_1(X)$ contains all monomials whose coefficients are coprime with p. One can easily prove by induction on i that

$$(A_1(X) + pA_2(X))^{p^i} \equiv (A_1(X))^{p^i} \pmod{p^{i+1}}.$$

We take $h = p^{k-1}$ so that $A^h(X) \equiv A_1^h(X) \pmod{p^k}$. It is easy to see that $f^{(h)}$ has form (11) with $b_{h\hat{L}} = a_{\hat{L}}^h$ and $b_{h\hat{R}} = a_{\hat{R}}^h$. □

Proof of Theorem 1. We prove the thesis only for the left Lyapunov exponent λ^- since the proof for λ^+ is analogous. We know that, since F is a linear map, Lyapunov exponents are independent of the particular configuration considered. Hence, in the rest of the proof we can safely write λ^- and Λ_n^- instead of $\lambda^-(x)$ and $\Lambda_n^-(x)$.

We first consider the case $m=p^k$ with p prime. From Lemma 1 we know that there exist $h\geqslant 1$ and $\hat{R}\in\mathbf{Z}$ such that $f^{(h)}$ is permutive in the variable $x_{h\hat{R}}$ and does not depend on any other variable x_j with $j>h\hat{R}$. Let $\lambda_{F^h}^-$ denote the left Lyapunov exponent of the map F^h. If $\hat{R}\leqslant 0$ we have that $f^{(h)}$ does not depend on variables with positive index. Hence, if two configurations differ in cells with index $i>\hat{\imath}$ such differences never "propagate" left under iteration of F. We conclude that $\lambda_{F^h}^-=0$. Assume now that $\hat{R}>0$. Let x and x' be two configurations such that $x(i)=x'(i)$ for every $i<\hat{\imath}$ and $x(\hat{\imath})\neq x'(\hat{\imath})$. Since $f^{(h)}$ is rightmost permutive, we have $[F^h(x)](i)=[F^h(x')](i)$ for every $i<\hat{\imath}-h\hat{R}$ and $x(\hat{\imath}-h\hat{R})\neq x'(\hat{\imath}-h\hat{R})$, i.e., the difference in cell $\hat{\imath}$ moves left of exactly $h\hat{R}$ positions. Hence $\lambda_{F^h}^-=h\hat{R}$. We now show that $\lambda_{F^h}^-=h\hat{R}$ implies $\lambda_F^-=\hat{R}$. From (6) we have

$$\lambda_F^- = \lim_{n\to\infty}\frac{1}{n}\Lambda_n^- = \lim_{n\to\infty}\frac{1}{nh}\Lambda_{nh}^- = \frac{1}{h}\lim_{n\to\infty}\frac{1}{n}\Lambda_{nh}^- = \frac{1}{h}\lambda_{F^h}^-.$$

Since $R=\max(0,\hat{R})$ then $\lambda_F^-=R$ and the thesis follows.

Consider now the general case $m=pq$, where $\gcd(p,q)=1$. By the Chinese Remainder Theorem we know that the ring $\mathbf{Z}_m$ is isomorphic to the direct product $\mathbf{Z}_p\otimes\mathbf{Z}_q$. Hence, F^n can be expressed as a linear combination of $[F]_p^n$ and $[F]_q^n$ as follows

$$F^n = \alpha q[F]_p^n + \beta p[F]_q^n$$

where $[\alpha q]_p=1$ and $[\beta p]_q=1$. This means that two configurations $F^n(x)$ and $F^n(y)$ differ in the ith cell if and only if the configurations $[F]_p^n([x]_p)$ and $[F]_p^n([y]_p)$ *or* the configurations $[F]_q^n([x]_q)$ and $[F]_q^n([y]_q)$ (or both) differ in the ith cell. Hence, differences in cells with index $i>\hat{\imath}$ can propagate left to the farthest cell reachable by either $[F]_p$ or $[F]_q$. Hence, $\lambda_F^-=\max(\lambda_{[F]_p}^-,\lambda_{[F]_q}^-)$. The thesis of the theorem follows by a simple inductive argument on the number of primes in the factorization of the modulus m. □

We now prove two lemmas which enable us to compute the entropy of linear CA.

Lemma 2. *Let F be a linear D-dimensional CA over $\mathscr{C}_m^D$ with $m=pq$ and* $\gcd(p,q)=1$. *Then*

$$\mathscr{H}(\mathscr{C}_{pq}^D,F) = \mathscr{H}(\mathscr{C}_p^D,[F]_p) + \mathscr{H}(\mathscr{C}_q^D,[F]_q).$$

Proof. Since $\gcd(p,q)=1$, the ring $\mathbf{Z}_m$ is isomorphic to the direct product $\mathbf{Z}_p\otimes\mathbf{Z}_q$. To prove the lemma we consider the characterization of the entropy given by (4). To each (hyper)rectangle R with elements in $\mathbf{Z}_m$ we can associate the pair of (hyper)rectangles R_p,R_q obtained by taking the content of each cell modulo p and modulo q. Since $\mathbf{Z}_m\simeq\mathbf{Z}_p\otimes\mathbf{Z}_q$ this is a bijective correspondence. In addition, R is generated through F by $x\in\mathscr{C}_m^D$, if and only if R_p (resp. R_q) is generated through $[F]_p$ (resp. $[F]_q$) by $[x]_p\in\mathscr{C}_q^D$ (resp. $[x]_q\in\mathscr{C}_q^D$). Hence, the number of (hyper)rectangles "realizable" by F

is equal to the product of the number of (hyper)rectangles "realizable" by $[F]_p$ and $[F]_q$ and the lemma follows. □

In view of the previous lemma, to compute the entropy of linear CA we can restrict our attention to linear CA defined over $\mathbf{Z}_{p^k}$ with p prime.

Lemma 3. *Let* $f(x_{-r},\dots,x_r)=[\sum_{i=-r}^{r} a_i x_i]_{p^k}$ *be any linear local rule defined over* $\mathbf{Z}_{p^k}$ *with* p *prime. Let* F *be the* 1*-dimensional global transition map associated to* f*. Let*

$$P = \{0\} \cup \{j:\ \gcd(a_j, p) = 1\}, \quad L = \min P, \quad \textit{and} \quad R = \max P.$$

Then

$$\mathscr{H}(\mathscr{C}^1_{p^k}, F) = k(R - L)\log(p).$$

Proof. From (7) and Theorem 1 we have

$$\mathscr{H}(\mathscr{C}^1_{p^k}, F) \leqslant k(R - L)\log(p). \tag{12}$$

Let $f^{(n)}$ be the local rule associated to $(\mathscr{C}^1_{p^k}, F^n)$. In general, $f^{(n)}$ has radius rn, i.e., it depends on at most $2rn+1$ variables. From Lemma 1 we have that there exist $h \geqslant 1$ and $\hat{L}, \hat{R} \in \mathbf{Z}$ such that F^h is permutive in the variables $x_{h\hat{L}}$ and $x_{h\hat{R}}$. In addition, F^h does not depend on variables x_j with $j < h\hat{L}$ or $j > h\hat{R}$. In other words, F^h is both leftmost and rightmost permutive. As a consequence $\mathscr{H}(\mathscr{C}^1_{p^k}, F^h)$ can be given as a function of $\hat{L}$ and $\hat{R}$ as follows. If $\hat{R} \geqslant \hat{L} \geqslant 0$ then $\mathscr{H}(\mathscr{C}^1_{p^k}, F^h) = hk\hat{R}\log(p)$, if $\hat{L} \leqslant \hat{R} \leqslant 0$ then $\mathscr{H}(\mathscr{C}^1_{p^k}, F^h) = -hk\hat{L}\log(p)$, and if both $\hat{L} < 0$ and $\hat{R} > 0$ then $\mathscr{H}(\mathscr{C}^1_{p^k}, F^h) = hk(\hat{R} - \hat{L})\log(p)$. The proof of these formulas is not difficult to obtain. Note that topological entropy of leftmost and rightmost permutive CA remains easy to compute also in the non-linear case. Since

$$L = \min(\hat{L}, 0) \quad \text{and} \quad R = \max(\hat{R}, 0), \tag{13}$$

we conclude that

$$\mathscr{H}(\mathscr{C}^1_{p^k}, F^h) = hk(R - L)\log(p). \tag{14}$$

From the definition of topological entropy we have

$$\begin{aligned}\mathscr{H}(\mathscr{C}^1_m, F) &= \lim_{w\to\infty}\left(\lim_{t\to\infty}\frac{\log R(w,t)}{t}\right) = \lim_{w\to\infty}\left(\lim_{t\to\infty}\frac{\log R(w,nt)}{nt}\right)\\ &\geqslant \frac{1}{n}\,\mathscr{H}(\mathscr{C}^1_m, F^n).\end{aligned} \tag{15}$$

From (14) and (15) we have

$$\mathscr{H}(\mathscr{C}^1_{p^k}, F) \geqslant \frac{\mathscr{H}(\mathscr{C}^1_{p^k}, F^h)}{h} = k(R - L)\log(p). \tag{16}$$

Combining (12) and (16) we obtain the thesis. □

Proof of Theorem 2. The proof easily follows from Lemma 3 and Corollary 4.2. □

Next result, which is a straightforward corollary of Theorems 1 and 2, characterizes the class of linear 1-dimensional CA $(\mathscr{C}_m^1, F)$ for which $\mathscr{H}(\mathscr{C}_m^1, F) = \mathscr{H}(\mathscr{C}_m^1, \sigma)(\lambda^+ + \lambda^-)$ where σ is the shift map defined in Example 1.

Corollary 4. *Let F be a 1-dimensional linear CA over $\mathbf{Z}_m$ based on the local rule f. Then $\mathscr{H}(\mathscr{C}_m^1, F) = \mathscr{H}(\mathscr{C}_m^1, \sigma)(\lambda^+ + \lambda^-)$ if and only if f is both leftmost and rightmost permutive.*

In order to compute the topological entropy for $D \geqslant 2$ we need a technical lemma which is the generalization of Lemma 1 to the case $D \geqslant 2$. Note that for $D \geqslant 2$ we do not have the concept of leftmost or rightmost permutivity, hence, we use a slightly different formulation. The common point between the two lemmas is that there exists $h \geqslant 1$ such that the coefficients of the local rule f which are multiple of p do not affect F^h.

Lemma 5. *Let $(\mathscr{C}_{p^k}^D, F)$ be a linear CA over $\mathbf{Z}_{p^k}$ (p prime) with local rule $f(x_1, \ldots, x_s) = [\sum_{i=1}^s \lambda_i x_i]_{p^k}$, and neighborhood vectors $\vec{u}_1, \ldots, \vec{u}_s$. Define*

$$I = \{i \mid \gcd(\lambda_i, p) = 1\}, \quad \hat{f} = \left[\sum_{i \in I} \lambda_i x_i\right]_{p^k},$$

and let $\hat{F}$ the global map associated to $\hat{f}$. Then, there exists $h \geqslant 1$ such that $F^h \equiv \hat{F}^h$, that is, $F^h(c) = \hat{F}^h(c)$ for any configuration $c \in \mathscr{C}_{p^k}^D$.

Proof. To the local map f we associate the formal power series

$$A(X_1, \ldots, X_D) = \sum_{i=1}^s \lambda_i X_1^{-\vec{u}_i(1)} \cdots X_D^{-\vec{u}_i(D)},$$

where $\vec{u}_i(j)$ denotes the jth component of vector $\vec{u}_i$. We know that the formal power series associated to $f^{(h)}$ is $(A(X_1, \ldots, X_D))^h$. The proof is obtained by taking $h = p^{k-1}$ and reasoning as in Lemma 1. □

Proof of Theorem 3. We know (see [12]) that a linear CA is either sensitive or equicontinuous. Since the entropy of any equicontinuous CA is zero (see for example [2]) we need only to prove that the entropy of a linear sensitive CA $(\mathscr{C}_m^D, F)$, $D \geqslant 2$, is unbounded. Let $\vec{u}_1, \ldots, \vec{u}_s$, $\lambda_1, \ldots, \lambda_s$ denote the neighborhood vectors and the corresponding coefficients of the map F (cf. (2)). By Theorem 5, we know that F is sensitive if and only if there exists a prime factor p of m such $p \nmid \lambda_i$ with $\vec{u}_i \neq \vec{0}$. Let k denote the power of p in the factorization of m. We now show that $\mathscr{H}(\mathscr{C}_{p^k}^D, [F]_{p^k}) = \infty$, which, by Lemma 2, proves the theorem. By Theorem 5, we get that $(\mathscr{C}_{p^k}^D, [F]_{p^k})$ is itself a

sensitive CA. Hence, it suffices to show that every D-dimensional ($D \geqslant 2$) sensitive CA over $\mathbf{Z}_{p^k}$ has unbounded entropy. For simplicity, we consider only the case $D=2$. For $D>2$ we only need to manage a more complex notation without introducing any new idea.

Let $(\mathscr{C}^2_{p^k}, F)$ denote a sensitive CA. We construct a set of 2-dimensional configurations whose evolutions according to F differentiate inside a space-temporal region of size $w \times w \times t$ which is the 2-dimensional version of the region depicted in Fig. 1. Then we prove that the cardinality of this set of configurations grows with w and t at a rate that makes the entropy unbounded. We proceed as follows. Let $\vec{u}_1, \ldots, \vec{u}_s$, $\lambda_1, \ldots, \lambda_s$ denote the neighborhood vectors and the corresponding coefficients of the map F. Let h be defined as in Lemma 5. Let $f^{(n)}$ be the local rule associated to $(\mathscr{C}^2_{p^k}, F^n)$. Let $\vec{u}(F)$ and $\lambda(F)$ be a neighborhood vector of F of maximum norm and the corresponding coefficient, respectively. From the sensitivity of $(\mathscr{C}^2_{p^k}, F)$ and from Lemma 5 we conclude that $\gcd(\lambda(F^h), p)=1$. Assume, without loss of generality, that $\rho = \vec{u}(F^h)(1) \geqslant \vec{u}(F^h)(2) \geqslant 0$.

We now show that given any set $\{z_{ij} \in \mathbf{Z}_{p^k} \colon i \in \mathbf{N},\ j \in \mathbf{Z}\}$ of elements of $\mathbf{Z}_{p^k}$ we can find a sequence $\{x_i \in \mathscr{C}^2_{p^k} \colon i \in \mathbf{N}\}$ of configurations such that

$$F^h(x_i) = x_{i+1} \quad \text{and} \quad x_i(0,j) = z_{ij} \ \ \forall j \in \mathbf{Z}. \tag{17}$$

In order to construct the above sequence we take advantage of the fact that the map F^h satisfies the property stated in Lemma 5 which is the extension of permutivity (defined for 1-dimensional CA) to the 2-dimensional case. We only give the basic steps of the construction procedure without going through the details.

We proceed as follows. Let $c_1 \in \mathscr{C}^2_{p^k}$ be any configuration such that $c_1(0,j)=z_{1j}$. We are not sure that $[F^h(c_1)](0,j)=z_{2j}$ for every $j \in \mathbf{Z}$. Since $\gcd((F^h), p^k)=1$ and $\vec{u}(F^h)(1) \geqslant \vec{u}(F^h)(2)$ we may find a configuration c_1'—obtained modifying c_1 at positions $(\rho h, j)$ with $j \in \mathbf{Z}$—for which $[F^h(c_1')](0,j)=z_{2j}$ for every $j \in \mathbf{Z}$. Set $c_2 = F^h(c_1')$. Again, we may find a configuration c_1''—obtained modifying c_1' at positions $(2\rho h, j)$ with $j \in \mathbf{Z}$—for which $[F^h(c_1'')](0,j)=z_{1j}$ and $[F^{2h}(c_1'')](0,j)=z_{2j}$ for every $j \in \mathbf{Z}$. Set $c_2' = F^h(c_1'')$ and $c_3 = F^{2h}(c_1'')$.

By iterating the above described procedure we construct sequences of configurations of type $S_i = c_i, c_i', c_i'', \ldots$, $i \geqslant 1$. For every $i \geqslant 1$ let $l_i \in \mathscr{C}^2_{p^k}$ be the limit of S_i which exists since $\mathscr{C}^2_{p^k}$ is a compact space. A simple but tedious calculation shows that $\{l_i \in \mathscr{C}^2_{p^k} \colon i \in \mathbf{N}\}$ satisfies (17).

We now show how to link the above constructed sequences (one for each set z_{ij}) of configurations to the computation of topological entropy. Let $Square(w,x) \in \mathbf{Z}_{p^k}^{w \times w}$ be the content of x at positions (i,j) with $-w < i \leqslant 0$ and $0 \leqslant j < w$. We have that $R(w^{(2)}, t)$ is equal to the number of distinct sequences $\langle Square(w,x), Square(w,F(x)), Square(w, F^2(x)), \ldots \rangle$ which can be obtained by varying $x \in \mathscr{C}^2_{p^k}$. It is not difficult to see that in view of (17) we can assign to the entries $(0,j)$, $0 \leqslant j < w$ of $Square(w, F^i(x))$ arbitrarily chosen elements of $\mathbf{Z}_{p^k}$ and still find a configuration x which realizes the sequence $\langle Square(w,x), Square(w,F(x)), Square(w,F^2(x)), \ldots \rangle$. Summarizing, we have

$$\bar{y}_s = \cdots y_s y_s y_s \overbrace{y_s \cdots y_s}^{l_t} y_s y_s y_s \cdots$$
$$\Downarrow$$
$$\hat{y}_s = \cdots y_s y_s y_s \underbrace{y_t \cdots y_t}_{l_s} y_s y_s y_s \cdots$$

Fig. 2. Construction of $\hat{y}_s$ given $\bar{y}_s$.

that $R(w^{(2)}, ht) \geqslant p^{kwt}$ and then

$$\mathscr{H}(\mathscr{C}^2_{p^k}, F) = \lim_{w\to\infty} \lim_{t\to\infty} \frac{\log R(w^{(2)}, t)}{t} = \lim_{w\to\infty} \lim_{t\to\infty} \frac{\log R(w^{(2)}, ht)}{ht}$$
$$\geqslant \lim_{w\to\infty} \lim_{t\to\infty} \frac{\log(p^{kwt})}{ht} = \lim_{w\to\infty} \frac{kw}{h} = \infty. \qquad \square$$

4.2. Positively expansive CA

We now prove Theorems 5 and 6 whose combination makes it possible to compute the topological entropy of positively expansive CA.

Proof of Theorem 5. Hedlund proved in [4] that open CA are μ-to-one mappings, i.e., the cardinality of the set of predecessors of any configuration is μ. Since positively expansive CA are open, we conclude that they are also μ-to-one mappings. Let $a \in \mathbf{Z}_m$. Since every configuration has the same number of predecessors, in order to evaluate the constant μ for the map F it suffices to determine the number of predecessors of the configuration $\bar{a} \in \mathscr{C}^1_m$, that is, the configuration such that $\bar{a}(i) = a$ for all $i \in \mathbf{Z}$. Since F is positively expansive it is surjective. We know that every predecessor of a spatially periodic configuration ($c \in \mathscr{C}^1_m$ is spatially periodic iff there exists $n \in \mathbf{N}$ such that $\sigma^n(c) = c$) according to a surjective CA is spatially periodic (see [3]). Thus, in order to count the number of predecessors of $\bar{a}$ it suffices to restrict our attention to spatially periodic configurations.

In what follows we prove that the number of spatially periodic predecessors of $\bar{a}$ is given by the sum of the lengths of all distinct cycles of the fp-graph $G_F = (V, E)$ (defined in Section 3.2) labeled by sequences of type a^n with $n \geqslant 1$. For any cycle

$$s = s_1 \xrightarrow{a} s_2 \xrightarrow{a} \cdots \xrightarrow{a} s_n \xrightarrow{a} s_1 \tag{18}$$

in the fp-graph G_F we define

$$y_s = s_1(1), s_2(1), s_3(1), \ldots, s_n(1), \tag{19}$$

where $s_i(j)$ denotes the jth character of the node $s_i \in V$. Consider now the configuration $\bar{y}_s$, that is, the configuration such that $\bar{y}_s(i) = s_{(i \bmod n)}(1)$. By construction we have that any sequence of $2r$ consecutive symbols in $\bar{y}_s$ is equal to a node in the cycle s. That is, $\bar{y}_s(1) \cdots \bar{y}_s(2r)$ is equal to s_1, $\bar{y}_s(2) \cdots \bar{y}_s(2r+1)$ is equal to s_2, and in general

$\bar{y}_s(i)\cdots\bar{y}_s(2r+i-1)$ is equal to $s_{(i \bmod n)}(1)$. Since each edge (s_i, s_{i+1}) is labeled with a we get that $F(\bar{y}_s)=\bar{a}$. In addition, the following properties hold.

(i) Every spatially periodic configuration $\bar{y}$ determines a cycle in the fp-graph G_F. Moreover, if $F(\bar{y})=\bar{a}$ all edges in the cycle are labeled by a. Thus, any predecessor of $\bar{a}$ corresponds to one of such cycles in the fp-graph.

(ii) All cycles with edges labeled with a are disjoint, i.e., they do not have common nodes (note that this implies that such cycles are also simple, i.e., all nodes in the cycle are distinct). Assume by contradiction that there exist two cycles s and t of length l_s and l_t which share (without loss of generality) their first node. Let $\hat{y}_s$ denote the configuration obtained from $\bar{y}_s$ replacing l_t copies of y_s with l_s copies of y_t (see Figure 2). We have that both $\hat{y}_s$ and $\bar{y}_s$ are predecessors of $\bar{a}$ according to F. However $\hat{y}_s$ and $\bar{y}_s$ differ in a finite number of positions which is impossible by Theorem 9.3 in [4].

(iii) Every cycle of length n determines exactly n distinct predecessors of $\bar{a}$. To see this, consider again the cycle s given by (18) and, for $i=1,\ldots,n$, define

$$y_{s,i} = s_i(1)s_{i+1}(1)\cdots s_{i-1}(1).$$

Clearly $F(\bar{y}_{s,i})=\bar{a}$. We need to prove that the $\bar{y}_{s,i}$ are distinct. It is easy to see that $y_{s,i}(1)\cdots y_{s,i}(2r)$ is equal to s_i. Since the nodes s_i are distinct the same is true for the n configurations $\bar{y}_{s,i}$.

The theorem is proven combining properties (i)–(iii). □

To prove Theorem 6 we need to recall the concept of topological conjugation. We say that two dynamical systems (X,F) and (X',F') are *topologically conjugated* if there exists a bijective function $\theta: X\to X'$ such that $\theta(F(x))=F'(\theta(x))$ and both θ and θ^{-1} are continuous (that is, θ is a homeomorphism between X and X'). If (X,F) and (X',F') are topologically conjugated then $\mathscr{H}(X,F)=\mathscr{H}(X',F')$.

Proof of Theorem 6. Since F is positively expansive then it is topologically conjugated to a suitable defined one-sided full shift σ defined over the alphabet $\mathscr{B}$ (see Theorem 8.5 in [15] and the footnote at the beginning of Section 3.2). As a consequence, $\mathscr{H}(\mathscr{C}_m^1,F)=\mathscr{H}(\mathscr{B}^{\mathbf{N}},\sigma)$. Since the entropy of any one-sided full shift is the logarithm of the number of its predecessors and topological conjugations preserve the number of predecessors, we conclude that $\mathscr{H}(\mathscr{C}_m^1,F)$ is equal to the logarithm of the number of the predecessors of F. □

Note that for non-expansive CA $(\mathscr{C}_m^1,F)$ the logarithm of the number of the predecessors of F is not related to its topological entropy. As an example the logarithm of the number of predecessors of the shift CA is zero while its entropy is equal to $\log m$.

5. Conclusions

This paper contains three main results:

- a formula for the Lyapunov exponents of linear CA (Theorem 1);

- a formula for the entropy of linear CA (Theorems 2 and 3);
- an algorithm for computing the entropy of positively expansive CA (Theorems 5 and 6).

The above results are a first step in the direction of understanding for which classes of CA Lyapunov exponents and topological entropy are computable. We are currently investigating the (un)decidability of topological properties such as (positive) expansivity, sensitivity, equicontinuity, and transitivity for general CA. In view of Theorem 3, an interesting open question is whether there exists a D-dimensional CA with $D \geqslant 2$ with finite non-zero topological entropy.

Acknowledgements

We would like to thank Masakazu Nasu for pointing out several important results in the field of symbolic dynamics.

References

[1] S. Amoroso, Y.N. Patt, Decision procedures for surjectivity and injectivity of parallel maps for tesselation structures, J. Comput. System Sci. 6 (1972) 448–464.
[2] F. Blanchard, P. Kurka, A. Maass, Topological and measure-theoretic properties of one-dimensional cellular automata, Physica D 103 (1997) 86–99.
[3] G. Cattaneo, E. Formenti, G. Manzini, L. Margara, Ergodicity, transitivity, and regularity for additive cellular automata over Z_m, Theoret. Comput. Sci. 233 (1–2) (2000) 147–164.
[4] G.A. Hedlund, Endomorphisms and automorphisms of the shift dynamical system, Math. Systems Theory 3 (1969) 320–375.
[5] L.P. Hurd, J. Kari, K. Culik, The topological entropy of cellular automata is uncomputable, Ergodic Theory Dynamical Systems 12 (1992) 255–265.
[6] M. Ito, N. Osato, M. Nasu, Linear cellular automata over Z_m, J. Comput. System Sci. 27 (1983) 125–140.
[7] J. Kari, The nilpotency problem of one-dimensional cellular automata, SIAM J. Comput. 21 (3) (1992) 571–586.
[8] J. Kari, Reversibility and surjectivity problems of cellular automata, J. Comput. System Sci. 48 (1) (1994) 149–182.
[9] P. Kurka, Languages, equicontinuity and attractors in cellular automata, Ergodic Theory Dynamical Systems 17 (1997) 417–433.
[10] G. Manzini, L. Margara, Invertible linear cellular automata over Z_m: Algorithmic and dynamical aspects, J. Comput. System Sci. 56 (1) (1998) 60–67.
[11] G. Manzini, L. Margara, Attractors of D-dimensional linear cellular automata, J. Comput. System Sci. 58 (3) (1999) 597–610.
[12] G. Manzini, L. Margara, A complete and efficiently computable topological classification of linear cellular automata over Z_m, Theoret. Comput. Sci. 221 (1999) 157–177.
[13] G. Morris, Dynamical constraints on group actions, Ph.D. Thesis, University of East Anglia, 1998.
[14] G. Morris, T. Ward, Entropy bounds for endomorphisms commuting with K actions, Israel J. Math. 106 (1998) 1–12.
[15] M. Nasu, Maps in symbolic dynamics, Lecture Notes of the 10th KAIST Mathematics Workshop, 1995.
[16] M. Nasu, Textile systems for endomorphisms and automorphisms of the shift, Mem. Amer. Math. Soc. 114 (546) (1995).
[17] M.A. Shereshevsky, Lyapunov exponents for one-dimensional cellular automata, J. Nonlinear Sci. 2 (1) (1992) 1.
[18] M.A. Shereshevsky, Expansiveness, entropy and polynomial growth for groups acting on subshifts by automorphisms, Indag. Math. 4 (1993) 203–210.